TRANSLATOR'S DECLARATION

I, Janet Hope, BSc(Hons.), MIL., MITI., translator to Messrs. Taylor and Meyer of 20 Kingsmead Road, London, SW2 3JD, Great Britain, verify that I know well both the German and the English language, that I have prepared the attached English translation of 33 pages of a German Patent application in the German language with the title:

Neue für das cstA-Gen kodierende Nukleotidsequenzen

identified by the code number 000173 BT at the upper left of each page and that the attached English translation of this document is a true and correct translation of the document attached thereto to the best of my knowledge and belief.

I further declare that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements are made with the knowledge that wilful false statements and the like are punishable by fine or imprisonment, or both, under 18 USC 1001, and that such false statements may jeopardize the validity of this document.

Signed: 

Dated: 13th November 2003

32301.195

FEDERAL REPUBLIC OF GERMANY

Certificate of Priority for Filing of a Patent Application

Filing number:	100 42 051.6
Filing date:	26th August 2000
Applicant/Proprietor:	Degussa AG, Düsseldorf/Germany First applicant: Degussa-Hüls AG, Frankfurt am Main/Germany
Title:	New nucleotide sequences which code for the cstA gene
IPC:	C 07 H, C 12 N, C 12 Q

The attached papers are a true and accurate reproduction of the original documents for this patent application.

Munich, 5th July 2001

On behalf of the President of the German Patent and Trade Mark Office

(signature)

Sieck

32301.195

New nucleotide sequences which code for the cstA gene

The invention provides nucleotide sequences from coryneform bacteria which code for the cstA gene and a process for the fermentative preparation of amino acids, in particular L-lysine, using bacteria in which the cstA gene is enhanced.

Prior art

L-Amino acids, in particular L-lysine, are used in human medicine and in the pharmaceuticals industry, in the foodstuffs industry and very particularly in animal nutrition.

It is known that amino acids are prepared by fermentation from strains of coryneform bacteria, in particular Corynebacterium glutamicum. Because of their great importance, work is constantly being undertaken to improve the preparation processes. Improvements to the process can relate to fermentation measures, such as, for example, stirring and supply of oxygen, or the composition of the nutrient media, such as, for example, the sugar concentration during the fermentation, or the working up to the product form by, for example, ion exchange chromatography, or the intrinsic output properties of the microorganism itself.

Methods of mutagenesis, selection and mutant selection are used to improve the output properties of these microorganisms. Strains which are resistant to antimetabolites or are auxotrophic for metabolites of regulatory importance and produce amino acids are obtained in this manner.

Methods of the recombinant DNA technique have also been employed for some years for improving the strain of Corynebacterium strains which produce L-amino acid, by amplifying individual amino acid biosynthesis genes and investigating the effect on the amino acid production.

Object of the invention

Where L-amino acids or amino acids are mentioned in the following, this means one or more amino acids, including their salts, chosen from the group consisting of L-asparagine, L-threonine, L-serine, L-glutamate, L-glycine, L-alanine, L-cysteine, L-valine, L-methionine, L-isoleucine, L-leucine, L-tyrosine, L-phenylalanine, L-histidine, L-lysine, L-tryptophan and L-arginine.

The inventors had the object of providing new measures for improved fermentative preparation of amino acids, in particular L-lysine.

Description of the invention

When lysine or L-lysine are mentioned in the following, not only the base but also the salts, such as e.g. lysine monohydrochloride or lysine sulfate, are meant by this.

The invention provides an isolated polynucleotide from coryneform bacteria, comprising a polynucleotide sequence which codes for the cstA gene, chosen from the group consisting of

a) polynucleotide which is identical to the extent of at least 70% to a polynucleotide which codes for a polypeptide which comprises the amino acid sequence of SEQ ID No. 2,

b) polynucleotide which codes for a polypeptide which comprises an amino acid sequence which is identical to the extent of at least 70% to the amino acid sequence of SEQ ID No. 2,

c) polynucleotide which is complementary to the polynucleotides of a) or b), and

d) polynucleotide comprising at least 15 successive nucleotides of the polynucleotide sequence of a), b) or c),

the polypeptide preferably having the activity of carbon starvation protein A.

The invention also provides the abovementioned polynucleotide, this preferably being a DNA which is capable of replication, comprising:

(i) the nucleotide sequence shown in SEQ ID No. 1, or

(ii) at least one sequence which corresponds to sequence (i) within the range of the degeneration of the genetic code, or

(iii) at least one sequence which hybridizes with the sequence complementary to sequence (i) or (ii), and optionally

(iv) sense mutations of neutral function in (i).

The invention also provides

a polynucleotide comprising the nucleotide sequence as shown in SEQ ID No. 1;

a polynucleotide which codes for a polypeptide which comprises the amino acid sequence as shown in SEQ ID No. 2;

a vector containing the polynucleotide according to the invention, in particular a shuttle vector or plasmid vector, and

coryneform bacteria serving as the host cell, which contain the vector or in which the cstA gene is enhanced.

The invention also provides polynucleotides which substantially comprise a polynucleotide sequence, which are

obtainable by screening by means of hybridization of a corresponding gene library of a coryneform bacterium, which comprises the complete gene or parts thereof, with a probe which comprises the sequence of the polynucleotide according to the invention according to SEQ ID No.1 or a fragment thereof, and isolation of the polynucleotide sequence mentioned.

Polynucleotide sequences according to the invention are suitable as hybridization probes for RNA, cDNA and DNA, in order to isolate, in the full length, nucleic acids or polynucleotides or genes which code for carbon starvation protein A, or to isolate those nucleic acids or polynucleotides or genes which have a high similarity of sequence with that of the cstA gene.

Polynucleotide sequences according to the invention are furthermore suitable as primers with the aid of which DNA of genes which code for carbon starvation protein A can be prepared with the polymerase chain reaction (PCR).

Such oligonucleotides which serve as probes or primers comprise at least 30, preferably at least 20, very particularly preferably at least 15 successive nucleotides. Oligonucleotides which have a length of at least 40 or 50 nucleotides are also suitable.

"Isolated" means separated out of its natural environment.

"Polynucleotide" in general relates to polyribonucleotides and polydeoxyribonucleotides, it being possible for these to be non-modified RNA or DNA or modified RNA or DNA.

The polynucleotides according to the invention include a polynucleotide according to SEQ ID No. 1 or a fragment prepared therefrom and also those which are at least 70%, preferably at least 80% and in particular at least 90% to 95% identical to the polynucleotide according to SEQ ID No. 1 or a fragment prepared therefrom.

"Polypeptides" are understood as meaning peptides or proteins which comprise two or more amino acids bonded via peptide bonds.

The polypeptides according to the invention include a polypeptide according to SEQ ID No. 2, in particular those with the biological activity of carbon starvation protein A, and also those which are at least 70%, preferably at least 80% and in particular which are at least 90% to 95% identical to the polypeptide according to SEQ ID No. 2 and have the activity mentioned.

The invention moreover provides a process for the fermentative preparation of amino acids, in particular L-lysine, using coryneform bacteria which in particular already produce amino acids, and in which the nucleotide sequences which code for the cstA gene are enhanced, in particular over-expressed.

The term "enhancement" in this connection describes the increase in the intracellular activity of one or more enzymes in a microorganism which are coded by the corresponding DNA, for example by increasing the number of copies of the gene or genes, using a potent promoter or using a gene which codes for a corresponding enzyme having a high activity, and optionally combining these measures.

The microorganisms which the present invention provides can prepare L-amino acids, in particular L-lysine, from glucose, sucrose, lactose, fructose, maltose, molasses, starch, cellulose or from glycerol and ethanol. They can be representatives of coryneform bacteria, in particular of the genus Corynebacterium. Of the genus Corynebacterium, there may be mentioned in particular the species Corynebacterium glutamicum, which is known among experts for its ability to produce L-amino acids.

Suitable strains of the genus Corynebacterium, in particular of the species Corynebacterium glutamicum (C. glutamicum), are in particular the known wild-type strains

Corynebacterium glutamicum ATCC13032
Corynebacterium acetoglutamicum ATCC15806
Corynebacterium acetoacidophilum ATCC13870
Corynebacterium thermoaminogenes FERM BP-1539
Corynebacterium melassecola ATCC17965
Brevibacterium flavum ATCC14067
Brevibacterium lactofermentum ATCC13869 and
Brevibacterium divaricatum ATCC14020

and L-lysine-producing mutants or strains prepared therefrom, such as, for example

Corynebacterium glutamicum FERM-P 1709
Brevibacterium flavum FERM-P 1708
Brevibacterium lactofermentum FERM-P 1712
Corynebacterium glutamicum FERM-P 6463
Corynebacterium glutamicum FERM-P 6464 and
Corynebacterium glutamicum DSM5715.

The inventors have succeeded in isolating the new cstA gene of C. glutamicum which codes for carbon starvation protein A.

To isolate the cstA gene or also other genes of C. glutamicum, a gene library of this microorganism is first set up in Escherichia coli (E. coli). The setting up of gene libraries is described in generally known textbooks and handbooks. The textbook by Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Germany, 1990), or the handbook by Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989) may be mentioned as an example. A well-known gene library is that of the E. coli K-12 strain W3110 set up in λ vectors by Kohara et al. (Cell 50, 495 -508 (1987)). Bathe et al. (Molecular and

General Genetics, 252:255-265, 1996) describe a gene library of C. glutamicum ATCC13032, which was set up with the aid of the cosmid vector SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84:2160-2164) in the E. coli K-12 strain NM554 (Raleigh et al., 1988, Nucleic Acids Research 16:1563-1575).

Börmann et al. (Molecular Microbiology 6(3), 317-326) (1992)) in turn describe a gene library of C. glutamicum ATCC13032 using the cosmid pHC79 (Hohn and Collins, Gene 11, 291-298 (1980)). To prepare a gene library of C. glutamicum in E. coli it is also possible to use plasmids such as pBR322 (Bolivar, Life Sciences, 25, 807-818 (1979)) or pUC9 (Vieira et al., 1982, Gene, 19:259-268). Suitable hosts are, in particular, those E. coli strains which are restriction- and recombination-defective. An example of these is the strain DH5αmcr, which has been described by Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649). The long DNA fragments cloned with the aid of cosmids can in turn be subcloned in the usual vectors suitable for sequencing and then sequenced, as is described e.g. by Sanger et al. (Proceedings of the National Academy of Sciences of the United States of America, 74:5463-5467, 1977).

The resulting DNA sequences can then be investigated with known algorithms or sequence analysis programs, such as e.g. that of Staden (Nucleic Acids Research 14, 217-232(1986)), that of Marck (Nucleic Acids Research 16, 1829-1836 (1988)) or the GCG program of Butler (Methods of Biochemical Analysis 39, 74-97 (1998)).

The new DNA sequence of C. glutamicum which codes for the cstA gene and which, as SEQ ID No. 1, is a constituent of the present invention has been obtained in this manner. The amino acid sequence of the corresponding protein has furthermore been derived from the present DNA sequence by

the methods described above. The resulting amino acid sequence of the cstA gene product is shown in SEQ ID No. 2.

Coding DNA sequences which result from SEQ ID No. 1 by the degeneracy of the genetic code are also a constituent of the invention. In the same way, DNA sequences which hybridize with SEQ ID No. 1 or parts of SEQ ID No. 1 are a constituent of the invention. Conservative amino acid exchanges, such as e.g. exchange of glycine for alanine or of aspartic acid for glutamic acid in proteins, are furthermore known among experts as "sense mutations" which do not lead to a fundamental change in the activity of the protein, i.e. are of neutral function. It is furthermore known that changes on the N and/or C terminus of a protein cannot substantially impair or can even stabilize the function thereof. Information in this context can be found by the expert, inter alia, in Ben-Bassat et al. (Journal of Bacteriology 169:751-757 (1987)), in O'Regan et al. (Gene 77:237-251 (1989)), in Sahin-Toth et al. (Protein Sciences 3:240-247 (1994)), in Hochuli et al. (Bio/Technology 6:1321-1325 (1988)) and in known textbooks of genetics and molecular biology. Amino acid sequences which result in a corresponding manner from SEQ ID No. 2 are also a constituent of the invention.

In the same way, DNA sequences which hybridize with SEQ ID No. 1 or parts of SEQ ID No. 1 are a constituent of the invention. Finally, DNA sequences which are prepared by the polymerase chain reaction (PCR) using primers which result from SEQ ID No. 1 are a constituent of the invention. Such oligonucleotides typically have a length of at least 15 nucleotides.

Instructions for identifying DNA sequences by means of hybridization can be found by the expert, inter alia, in the handbook "The DIG System Users Guide for Filter Hybridization" from Boehringer Mannheim GmbH (Mannheim, Germany, 1993) and in Liebl et al. (International Journal

of Systematic Bacteriology (1991) 41: 255-260). The hybridization takes place under stringent conditions, that is to say only hybrids in which the probe and target sequence, i.e. the polynucleotides treated with the probe, are at least 70% identical are formed. It is known that the stringency of the hybridization, including the washing steps, is influenced or determined by varying the buffer composition, the temperature and the salt concentration. The hybridization reaction is preferably carried out under a relatively low stringency compared with the washing steps (Hybaid Hybridisation Guide, Hybaid Limited, Teddington, UK, 1996).

A 5x SSC buffer at a temperature of approx. 50 - 68°C, for example, can be employed for the hybridization reaction. Probes can also hybridize here with polynucleotides which are less than 70% identical to the sequence of the probe. Such hybrids are less stable and are removed by washing under stringent conditions. This can be achieved, for example, by lowering the salt concentration to 2x SSC and subsequently 0.5x SSC (The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim, Mannheim, Germany, 1995) a temperature of approx. 50 - 68°C being established. It is optionally possible to lower the salt concentration to 0.1x SSC. Polynucleotide fragments which are, for example, at least 70% or at least 80% or at least 90% to 95% identical to the sequence of the probe employed can be isolated by increasing the hybridization temperature stepwise in steps of approx. 1 - 2°C. Further instructions on hybridization are obtainable on the market in the form of so-called kits (e.g. DIG Easy Hyb from Roche Diagnostics GmbH, Mannheim, Germany, Catalogue No. 1603558).

Instructions for amplification of DNA sequences with the aid of the polymerase chain reaction (PCR) can be found by the expert, inter alia, in the handbook by Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press,

Oxford, UK, 1984) and in Newton and Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Germany, 1994).

It has been found that coryneform bacteria produce amino acids, in particular L-lysine, in an improved manner after over-expression of the cstA gene.

To achieve an over-expression, the number of copies of the corresponding genes can be increased, or the promoter and regulation region or the ribosome binding site upstream of the structural gene can be mutated. Expression cassettes which are incorporated upstream of the structural gene act in the same way. By inducible promoters, it is additionally possible to increase the expression in the course of fermentative lysine production. The expression is likewise improved by measures to prolong the life of the m-RNA. Furthermore, the enzyme activity is also increased by preventing the degradation of the enzyme protein. The genes or gene constructs can either be present in plasmids with a varying number of copies, or can be integrated and amplified in the chromosome. Alternatively, an over-expression of the genes in question can furthermore be achieved by changing the composition of the media and the culture procedure.

Instructions in this context can be found by the expert, inter alia, in Martin et al. (Bio/Technology 5, 137-146 (1987)), in Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya and Morinaga (Bio/Technology 6, 428-430 (1988)), in Eikmanns et al. (Gene 102, 93-98 (1991)), in European Patent Specification 0 472 869, in US Patent 4,601,893, in Schwarzer and Pühler (Bio/Technology 9, 84-87 (1991), in Reinscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)), in LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in Patent Application WO 96/15246, in Malumbres et al. (Gene 134, 15 - 24 (1993)), in Japanese Laid-Open Specification JP-A-10-229891, in Jensen and Hammer (Biotechnology and

Bioengineering 58, 191-195 (1998)), in Makrides (Microbiological Reviews 60:512-538 (1996)) and in known textbooks of genetics and molecular biology.

By way of example, for enhancement the cstA gene according to the invention was over-expressed with the aid of episomal plasmids. Suitable plasmids are those which are replicated in coryneform bacteria. Numerous known plasmid vectors, such as e.g. pZ1 (Menkel et al., Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns et al., Gene 102:93-98 (1991)) or pHS2-1 (Sonnen et al., Gene 107:69-74 (1991)) are based on the cryptic plasmids pHM1519, pBL1 or pGA1. Other plasmid vectors, such as e.g. those based on pCG4 (US-A 4,489,160), or pNG2 (Serwold-Davis et al., FEMS Microbiology Letters 66, 119-124 (1990)), or pAG1 (US-A 5,158,891), can be used in the same manner.

Plasmid vectors which are furthermore suitable are also those with the aid of which the process of gene amplification by integration into the chromosome can be used, as has been described, for example, by Reinscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)) for duplication or amplification of the hom-thrB operon. In this method, the complete gene is cloned in a plasmid vector which can replicate in a host (typically E. coli), but not in C. glutamicum. Possible vectors are, for example, pSUP301 (Simon et al., Bio/Technology 1, 784-791 (1983)), pK18mob or pK19mob (Schäfer et al., Gene 145, 69-73 (1994)), pGEM-T (Promega corporation, Madison, WI, USA), pCR2.1-TOPO (Shuman (1994). Journal of Biological Chemistry 269:32678-84; US-A 5,487,993), pCR®Blunt (Invitrogen, Groningen, Holland; Bernard et al., Journal of Molecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf et al, 1991, Journal of Bacteriology 173:4510-4516) or pBGS8 (Spratt et al.,1986, Gene 41: 337-342). The plasmid vector which contains the gene to be amplified is then transferred into the desired strain of C. glutamicum by conjugation or

transformation. The method of conjugation is described, for example, by Schäfer et al. (Applied and Environmental Microbiology 60, 756-759 (1994)). Methods for transformation are described, for example, by Thierbach et al. (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican and Shivnan (Bio/Technology 7, 1067-1070 (1989)) and Tauch et al. (FEMS Microbiological Letters 123, 343-347 (1994)). After homologous recombination by means of a "cross over" event, the resulting strain contains at least two copies of the gene in question.

In addition, it may be advantageous for the production of amino acids, in particular L-lysine, to enhance one or more enzymes of the particular biosynthesis pathway, of glycolysis, of anaplerosis, of the citric acid cycle or of amino acid export and optionally regulatory proteins, in addition to the cstA gene.

Thus, for example, for the preparation of amino acids, in particular L-lysine, one or more genes chosen from the group consisting of

- the dapA gene which codes for dihydrodipicolinate synthase (EP-B 0 197 335),
- the gap gene which codes for glyceraldehyde 3-phosphate dehydrogenase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the tpi gene which codes for triose phosphate isomerase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the pgk gene which codes for 3-phosphoglycerate kinase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the pyc gene which codes for pyruvate carboxylase (Peters-Wendisch et al.(Microbiology 144, 915 - 927 (1998)),

- the lysC gene which codes for a feed back resistant aspartate kinase (Kalinowski et al. (1990), Molecular and General Genetics 224, 317-324; Accession No.P26512),
- the lysE gene which codes for lysine export (DE-A-195 48 222),
- the zwa1 gene which codes for the Zwa1 protein (DE: 19959328.0, DSM 13115)

can be enhanced, in particular over-expressed.

It may furthermore be advantageous for the production of amino acids, in particular L-lysine, in addition to the enhancement of the cstA gene, for one or more genes chosen from the group consisting of

- the pck gene which codes for phosphoenol pyruvate carboxykinase (DE 199 50 409.1; DSM 13047),
- the pgi gene which codes for glucose 6-phosphate isomerase (US 09/396,478; DSM 12969),
- the poxB gene which codes for pyruvate oxidase (DE: 1995 1975.7; DSM 13114),
- the zwa2 gene which codes for the Zwa2 protein (DE: 19959327.2, DSM 13113)

to be attenuated, in particular for the expression thereof to be reduced.

In addition to over-expression of the cstA gene it may furthermore be advantageous, for the production of amino acids, in particular L-lysine, to eliminate undesirable side reactions, (Nakayama: "Breeding of Amino Acid Producing Micro-organisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).

The microorganisms prepared according to the invention can be cultured continuously or discontinuously in the batch process (batch culture) or in the fed batch (feed process) or repeated fed batch process (repetitive feed process) for the purpose of production of amino acids, in particular L-lysine. A summary of known culture methods are described in the textbook by Chmiel (Bioprozesstechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) or in the textbook by Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)).

The culture medium to be used must meet the requirements of the particular strains in a suitable manner. Descriptions of culture media for various microorganisms are contained in the handbook "Manual of Methods for General Bacteriology" of the American Society for Bacteriology (Washington D.C., USA, 1981).

Sugars and carbohydrates, such as e.g. glucose, sucrose, lactose, fructose, maltose, molasses, starch and cellulose, oils and fats, such as e.g. soya oil, sunflower oil, groundnut oil and coconut fat, fatty acids, such as e.g. palmitic acid, stearic acid and linoleic acid, alcohols, such as e.g. glycerol and ethanol, and organic acids, such as e.g. acetic acid, can be used as the source of carbon. These substance can be used individually or as a mixture.

Organic nitrogen-containing compounds, such as peptones, yeast extract, meat extract, malt extract, corn steep liquor, soya bean flour and urea, or inorganic compounds, such as ammonium sulfate, ammonium chloride, ammonium phosphate, ammonium carbonate and ammonium nitrate, can be used as the source of nitrogen. The sources of nitrogen can be used individually or as a mixture.

Phosphoric acid, potassium dihydrogen phosphate or dipotassium hydrogen phosphate or the corresponding sodium-containing salts can be used as the source of phosphorus.

The culture medium must furthermore comprise salts of metals, such as e. g. magnesium sulfate or iron sulfate, which are necessary for growth. Finally, essential growth substances, such as amino acids and vitamins, can be employed in addition to the abovementioned substances. Suitable precursors can moreover be added to the culture medium. The starting substances mentioned can be added to the culture in the form of a single batch, or can be fed in during the culture in a suitable manner.

Basic compounds, such as sodium hydroxide, potassium hydroxide, ammonia or aqueous ammonia, or acid compounds, such as phosphoric acid or sulfuric acid, can be employed in a suitable manner to control the pH of the culture. Antifoams, such as e.g. fatty acid polyglycol esters, can be employed to control the development of foam. Suitable substances having a selective action, such as e.g. antibiotics, can be added to the medium to maintain the stability of plasmids. To maintain aerobic conditions, oxygen or oxygen-containing gas mixtures, such as e.g. air, are introduced into the culture. The temperature of the culture is usually 20°C to 45°C, and preferably 25°C to 40°C. Culturing is continued until a maximum of the desired product has formed. This target is usually reached within 10 hours to 160 hours.

The analysis of lysine can be carried out by ion exchange chromatography with subsequent ninhydrin derivatization, as described by Spackman et al. (Analytical Chemistry, 30, (1958), 1190).

The process according to the invention is used for the fermentative preparation of amino acids, in particular L-lysine.

The present invention is explained in more detail in the following with the aid of embodiment examples.

The isolation of plasmid DNA from Escherichia coli and all techniques of restriction, Klenow and alkaline phosphatase treatment were carried out by the method of Sambrook et al. (Molecular Cloning. A Laboratory Manual (1989) Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY, USA). Methods for transformation of Escherichia coli are also described in this handbook.

The composition of the usual nutrient media, such as LB or TY medium, can also be found in the handbook by Sambrook et al.

Example 1

Preparation of a genomic cosmid gene library from Corynebacterium glutamicum ATCC 13032

Chromosomal DNA from Corynebacterium glutamicum ATCC 13032 was isolated as described by Tauch et al. (1995, Plasmid 33:168-179) and partly cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, Product Description Sau3AI, Code no. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Diagnostics GmbH, Mannheim, Germany, Product Description SAP, Code no. 1758250). The DNA of the cosmid vector SuperCos1 (Wahl et al. (1987) Proceedings of the National Academy of Sciences USA 84:2160-2164), obtained from Stratagene (La Jolla, USA, Product Description SuperCos1 Cosmid Vector Kit, Code no. 251301) was cleaved with the restriction enzyme XbaI (Amersham Pharmacia, Freiburg, Germany, Product Description XbaI, Code no. 27-0948-02) and likewise dephosphorylated with shrimp alkaline phosphatase.

The cosmid DNA was then cleaved with the restriction enzyme BamHI (Amersham Pharmacia, Freiburg, Germany, Product Description BamHI, Code no. 27-0868-04). The cosmid DNA treated in this manner was mixed with the treated ATCC13032 DNA and the batch was treated with T4 DNA ligase (Amersham

Pharmacia, Freiburg, Germany, Product Description T4-DNA-Ligase, Code no.27-0870-04). The ligation mixture was then packed in phages with the aid of Gigapack II XL Packing Extract (Stratagene, La Jolla, USA, Product Description Gigapack II XL Packing Extract, Code no. 200217).

For infection of the E. coli strain NM554 (Raleigh et al. 1988, Nucleic Acid Research 16:1563-1575) the cells were taken up in 10 mM $MgSO_4$ and mixed with an aliquot of the phage suspension. The infection and titering of the cosmid library were carried out as described by Sambrook et al. (1989, Molecular Cloning: A laboratory Manual, Cold Spring Harbor), the cells being plated out on LB agar (Lennox, 1955, Virology, 1:190) with 100 mg/l ampicillin. After incubation overnight at 37°C, recombinant individual clones were selected.

Example 2

Isolation and sequencing of the cstA gene

The cosmid DNA of an individual colony was isolated with the Qiaprep Spin Miniprep Kit (Product No. 27106, Qiagen, Hilden, Germany) in accordance with the manufacturer's instructions and partly cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, Product Description Sau3AI, Product No. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Diagnostics GmbH, Mannheim, Germany, Product Description SAP, Product No. 1758250). After separation by gel electrophoresis, the cosmid fragments in the size range of 1500 to 2000 bp were isolated with the QiaExII Gel Extraction Kit (Product No. 20021, Qiagen, Hilden, Germany).

The DNA of the sequencing vector pZero-1, obtained from Invitrogen (Groningen, The Netherlands, Product Description Zero Background Cloning Kit, Product No. K2500-01) was cleaved with the restriction enzyme BamHI (Amersham

Pharmacia, Freiburg, Germany, Product Description BamHI, Product No. 27-0868-04). The ligation of the cosmid fragments in the sequencing vector pZero-1 was carried out as described by Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor), the DNA mixture being incubated overnight with T4 ligase (Pharmacia Biotech, Freiburg, Germany). This ligation mixture was then electroporated (Tauch et al. 1994, FEMS Microbiol Letters, 123:343-7) into the E. coli strain DH5αMCR (Grant, 1990, Proceedings of the National Academy of Sciences U.S.A., 87:4645-4649) and plated out on LB agar (Lennox, 1955, Virology, 1:190) with 50 mg/l zeocin.

The plasmid preparation of the recombinant clones was carried out with Biorobot 9600 (Product No. 900200, Qiagen, Hilden, Germany). The sequencing was carried out by the dideoxy chain termination method of Sanger et al. (1977, Proceedings of the National Academy of Sciences U.S.A., 74:5463-5467) with modifications according to Zimmermann et al. (1990, Nucleic Acids Research, 18:1067). The "RR dRhodamin Terminator Cycle Sequencing Kit" from PE Applied Biosystems (Product No. 403044, Weiterstadt, Germany) was used. The separation by gel electrophoresis and analysis of the sequencing reaction were carried out in a "Rotiphoresis NF Acrylamide/Bisacrylamide" Gel (29:1) (Product No. A124.1, Roth, Karlsruhe, Germany) with the "ABI Prism 377" sequencer from PE Applied Biosystems (Weiterstadt, Germany).

The raw sequence data obtained were then processed using the Staden program package (1986, Nucleic Acids Research, 14:217-231) version 97-0. The individual sequences of the pZero1 derivatives were assembled to a continuous contig. The computer-assisted coding region analysis was prepared with the XNIP program (Staden, 1986, Nucleic Acids Research, 14:217-231).

The resulting nucleotide sequence is shown in SEQ ID No. 1. Analysis of the nucleotide sequence showed an open reading frame of 2316 base pairs, which was called the cstA gene. The cstA gene codes for a protein of 772 amino acids.

SEQUENCE PROTOCOL

<110> Degussa-Hüls AG

<120> New nucleotide sequences which code for the cstA gene

<130> 000173 BT

<140>
<141>

<160> 2

<170> PatentIn Ver. 2.1

<210> 1
<211> 2718
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (200)..(2515)
<223> cstA gene

<400> 1
aggatggtat aaatcatctc tcaatgttac ttttccattg ttaagaatta acaactctcg 60

gtgatttgtc gcatacccag ctgtcaaaga tccgatcatc ggcatacaga aacacccatc 120

tggccgaact ttccttttc tgcatgcatt tctgcacaca gtttctgccc gctgtttctg 180

cccgctgttt ctacgcata gtg gct ttg aaa cga ccc gaa gag aaa aca gta 232
Val Ala Leu Lys Arg Pro Glu Glu Lys Thr Val
1 5 10

aag atc gtg acc ata aaa cag act gac aac atc aat gac gat gat ttg 280
Lys Ile Val Thr Ile Lys Gln Thr Asp Asn Ile Asn Asp Asp Asp Leu
15 20 25

gtg tac agc aac gct act gac ctt cca gta ggc gtg aag aag tcc cct 328
Val Tyr Ser Asn Ala Thr Asp Leu Pro Val Gly Val Lys Lys Ser Pro
30 35 40

aaa atg tca ccg acc gcc cgc gtt ggt ctc ctt gtc ttt ggg gtt atc 376
Lys Met Ser Pro Thr Ala Arg Val Gly Leu Leu Val Phe Gly Val Ile
45 50 55

gcg gcg gtg ggt tgg gga gca atc gct ttc tcc cgt ggc gaa aca atc 424
Ala Ala Val Gly Trp Gly Ala Ile Ala Phe Ser Arg Gly Glu Thr Ile
60 65 70 75

aac tct gtg tgg ctg gtt ttg gcg gca gtt ggt tcc tat atc att gcg 472
Asn Ser Val Trp Leu Val Leu Ala Ala Val Gly Ser Tyr Ile Ile Ala
80 85 90

```
ttt tct ttc tat gcc cga ctg att gaa tac aaa gtt gtt aag ccg aaa   520
Phe Ser Phe Tyr Ala Arg Leu Ile Glu Tyr Lys Val Val Lys Pro Lys
            95                  100                 105

gat cag cga gca acc ccg gcg gaa tac gtt aat gac ggc aag gac tat   568
Asp Gln Arg Ala Thr Pro Ala Glu Tyr Val Asn Asp Gly Lys Asp Tyr
        110                 115                 120

gtc cca acg gat cgt cgt gtg ctt ttt ggc cac cac ttt gca gct att   616
Val Pro Thr Asp Arg Arg Val Leu Phe Gly His His Phe Ala Ala Ile
    125                 130                 135

gca ggt gcc ggt cca ttg gtt gga cct gtc atg gcc gcg cag atg ggc   664
Ala Gly Ala Gly Pro Leu Val Gly Pro Val Met Ala Ala Gln Met Gly
140                 145                 150                 155

tac ctg cca ggc acc ttg tgg att atc ctc ggt gtg att ttc gcc ggt   712
Tyr Leu Pro Gly Thr Leu Trp Ile Ile Leu Gly Val Ile Phe Ala Gly
                160                 165                 170

gca gtg cag gac tac cta gtg ctg tgg gtg tct act cgt agg cgt gga   760
Ala Val Gln Asp Tyr Leu Val Leu Trp Val Ser Thr Arg Arg Arg Gly
            175                 180                 185

cgc tca ctt ggc cag atg gtt cgt gat gaa atg ggc acg gtc ggt gga   808
Arg Ser Leu Gly Gln Met Val Arg Asp Glu Met Gly Thr Val Gly Gly
        190                 195                 200

gct gcc ggt atc ttg gcg acc atc tcc atc atg atc atc att atc gcg   856
Ala Ala Gly Ile Leu Ala Thr Ile Ser Ile Met Ile Ile Ile Ile Ala
    205                 210                 215

gtg ctc gca ttg atc gtg gtt aat gca ctg gct gat tca cca tgg ggc   904
Val Leu Ala Leu Ile Val Val Asn Ala Leu Ala Asp Ser Pro Trp Gly
220                 225                 230                 235

gtt ttc tcc atc acc atg acc atc cca att gca ctg ttc atg ggt gtg   952
Val Phe Ser Ile Thr Met Thr Ile Pro Ile Ala Leu Phe Met Gly Val
                240                 245                 250

tac ttg cgt tac ctg cgc cca ggt cgt gtt act gaa gtg tcc atc atc   1000
Tyr Leu Arg Tyr Leu Arg Pro Gly Arg Val Thr Glu Val Ser Ile Ile
            255                 260                 265

ggt gtg gca ctg ctc ctg ctg gct atc gtt gct ggt ggt tgg gtt gca   1048
Gly Val Ala Leu Leu Leu Leu Ala Ile Val Ala Gly Gly Trp Val Ala
        270                 275                 280

gac acc tca tgg ggc gtg gaa tgg ttc acc tgg tct aag acc act ttg   1096
Asp Thr Ser Trp Gly Val Glu Trp Phe Thr Trp Ser Lys Thr Thr Leu
    285                 290                 295

gcg ttg gcc ttg atc ggt tac gga atc atg gct gcg att ttg ccg gtg   1144
Ala Leu Ala Leu Ile Gly Tyr Gly Ile Met Ala Ala Ile Leu Pro Val
300                 305                 310                 315

tgg ctg ctg ctt gca ccg cgc gat tac ctg tct acc ttt atg aag atc   1192
Trp Leu Leu Leu Ala Pro Arg Asp Tyr Leu Ser Thr Phe Met Lys Ile
                320                 325                 330
```

```
ggc gtc atc ggt ctg ttg gca gtg ggt att ttg ttc gca cgt cct gag   1240
Gly Val Ile Gly Leu Leu Ala Val Gly Ile Leu Phe Ala Arg Pro Glu
            335                 340                 345

gtg cag atg cct tcc gtg acc tcc ttc gca ctt gag ggc aac ggt ccg   1288
Val Gln Met Pro Ser Val Thr Ser Phe Ala Leu Glu Gly Asn Gly Pro
        350                 355                 360

gtg ttc tct gga agt ctg ttc cca ttc ctg ttc atc acg att gcc tgt   1336
Val Phe Ser Gly Ser Leu Phe Pro Phe Leu Phe Ile Thr Ile Ala Cys
    365                 370                 375

ggt gca ctg tct ggt ttc cac gca ctg att tct tca gga acc aca cca   1384
Gly Ala Leu Ser Gly Phe His Ala Leu Ile Ser Ser Gly Thr Thr Pro
380                 385                 390                 395

aag ctt gtg gag aag gaa tcc cag atg cgc atg ctc ggc tac ggc ggc   1432
Lys Leu Val Glu Lys Glu Ser Gln Met Arg Met Leu Gly Tyr Gly Gly
                400                 405                 410

atg ttg atg gaa tct ttc gtg gcg atg atg gca ctg atc acc gct gtt   1480
Met Leu Met Glu Ser Phe Val Ala Met Met Ala Leu Ile Thr Ala Val
            415                 420                 425

att ctg gat cgt cac ctg tac ttc tcc atg aac gct ccg ctg gca ctg   1528
Ile Leu Asp Arg His Leu Tyr Phe Ser Met Asn Ala Pro Leu Ala Leu
        430                 435                 440

act ggt gga gat cca gca acc gca gct gag tgg gtt aac tcc att ggg   1576
Thr Gly Gly Asp Pro Ala Thr Ala Ala Glu Trp Val Asn Ser Ile Gly
    445                 450                 455

ctg aca ggt gcg gat atc acc ccg gaa cag ctg tcg gaa gct gct gaa   1624
Leu Thr Gly Ala Asp Ile Thr Pro Glu Gln Leu Ser Glu Ala Ala Glu
460                 465                 470                 475

agt gtc gga gaa tcc act gtt att tcc cgt acc ggt ggc gca cca acc   1672
Ser Val Gly Glu Ser Thr Val Ile Ser Arg Thr Gly Gly Ala Pro Thr
                480                 485                 490

ttg gcg ttc ggt atg tct gaa atc ctc tcc gga ttc atc ggc ggc gct   1720
Leu Ala Phe Gly Met Ser Glu Ile Leu Ser Gly Phe Ile Gly Gly Ala
            495                 500                 505

gga atg aag gcg ttc tgg tac cac ttc gcc atc atg ttt gag gct ctg   1768
Gly Met Lys Ala Phe Trp Tyr His Phe Ala Ile Met Phe Glu Ala Leu
        510                 515                 520

ttc atc ctc act act gtg gat gca ggt act cgt gtg gct cgc ttt atg   1816
Phe Ile Leu Thr Thr Val Asp Ala Gly Thr Arg Val Ala Arg Phe Met
    525                 530                 535

atg acc gat acc ttg ggc aat gtt cca ggt ctg cgc cgt ttc aag gat   1864
Met Thr Asp Thr Leu Gly Asn Val Pro Gly Leu Arg Arg Phe Lys Asp
540                 545                 550                 555
```

```
cct tca tgg act gtc ggt aac tgg att tct acc gtg ttt gtg tgt gct   1912
Pro Ser Trp Thr Val Gly Asn Trp Ile Ser Thr Val Phe Val Cys Ala
                560                 565                 570

cta tgg ggt gct att ttg ctc atg ggt gtt acc gat cca ctg ggc ggc   1960
Leu Trp Gly Ala Ile Leu Leu Met Gly Val Thr Asp Pro Leu Gly Gly
            575                 580                 585

atc aac gtg ctt ttc cca cta ttc ggt atc gct aac cag ctg ctc gcc   2008
Ile Asn Val Leu Phe Pro Leu Phe Gly Ile Ala Asn Gln Leu Leu Ala
        590                 595                 600

gct att gca ctt gct ctc gtg ctg gtt gtt gtg gtg aag aag ggc ctg   2056
Ala Ile Ala Leu Ala Leu Val Leu Val Val Val Lys Lys Gly Leu
    605                 610                 615

tac aag tgg gcg tgg att cca gct gtt cct ttg gca tgg gat ctc att   2104
Tyr Lys Trp Ala Trp Ile Pro Ala Val Pro Leu Ala Trp Asp Leu Ile
620                 625                 630                 635

gtc acg atg act gcg tca tgg cag aag att ttc cac tct gat ccg gct   2152
Val Thr Met Thr Ala Ser Trp Gln Lys Ile Phe His Ser Asp Pro Ala
                640                 645                 650

att ggc tac tgg gct cag aac gcg aac ttc cgc gat gca aag tct caa   2200
Ile Gly Tyr Trp Ala Gln Asn Ala Asn Phe Arg Asp Ala Lys Ser Gln
            655                 660                 665

ggc ctt acc gaa ttt ggt gcc gct aaa tct cct gag gca atc gat gcg   2248
Gly Leu Thr Glu Phe Gly Ala Ala Lys Ser Pro Glu Ala Ile Asp Ala
        670                 675                 680

gtt atc cga aac acc atg att cag ggc atc ttg tcc atc ctg ttc gcg   2296
Val Ile Arg Asn Thr Met Ile Gln Gly Ile Leu Ser Ile Leu Phe Ala
    685                 690                 695

gtg ctc gtc ctc gtt gtt gtc ggc gca gcc att gcg gtg tgc atc aag   2344
Val Leu Val Leu Val Val Val Gly Ala Ala Ile Ala Val Cys Ile Lys
700                 705                 710                 715

tcc atc agg gct cgt gca gcc gga aca cct ttg gag acc act gaa gag   2392
Ser Ile Arg Ala Arg Ala Ala Gly Thr Pro Leu Glu Thr Thr Glu Glu
                720                 725                 730

cct gat act gaa tct gag ttc ttc gcc cca act gga ttc ctt gca tct   2440
Pro Asp Thr Glu Ser Glu Phe Phe Ala Pro Thr Gly Phe Leu Ala Ser
            735                 740                 745

tcc agg gat aag gaa gtc cag gcc atg tgg gac gag cgc tac cca ggc   2488
Ser Arg Asp Lys Glu Val Gln Ala Met Trp Asp Glu Arg Tyr Pro Gly
        750                 755                 760

ggt gcg ccc gtg tct tct gga ggg cac taaaacatga tggctcttac         2535
Gly Ala Pro Val Ser Ser Gly Gly His
    765                 770

tcatgcactg tggaaaatcc cgcgggcggt gtggtggtat ctcactgagc tcatggggga 2595

cacggcgtat tccaagtatg tggtgcactt aaagcaccac catccggatg ctccgattcc 2655
```

tactgagcgg gagtattggc gggcaaagta tgcagatcag gacgctaatc ctggtgcccg 2715

ctg 2718

<210> 2
<211> 772
<212> PRT
<213> Corynebacterium glutamicum

<400> 2

Val	Ala	Leu	Lys	Arg	Pro	Glu	Glu	Lys	Thr	Val	Lys	Ile	Val	Thr	Ile
1				5				10						15	
Lys	Gln	Thr	Asp	Asn	Ile	Asn	Asp	Asp	Asp	Leu	Val	Tyr	Ser	Asn	Ala
			20					25					30		
Thr	Asp	Leu	Pro	Val	Gly	Val	Lys	Lys	Ser	Pro	Lys	Met	Ser	Pro	Thr
		35					40					45			
Ala	Arg	Val	Gly	Leu	Leu	Val	Phe	Gly	Val	Ile	Ala	Ala	Val	Gly	Trp
	50					55					60				
Gly	Ala	Ile	Ala	Phe	Ser	Arg	Gly	Glu	Thr	Ile	Asn	Ser	Val	Trp	Leu
65					70					75					80
Val	Leu	Ala	Ala	Val	Gly	Ser	Tyr	Ile	Ile	Ala	Phe	Ser	Phe	Tyr	Ala
				85					90					95	
Arg	Leu	Ile	Glu	Tyr	Lys	Val	Val	Lys	Pro	Lys	Asp	Gln	Arg	Ala	Thr
			100					105					110		
Pro	Ala	Glu	Tyr	Val	Asn	Asp	Gly	Lys	Asp	Tyr	Val	Pro	Thr	Asp	Arg
		115					120					125			
Arg	Val	Leu	Phe	Gly	His	His	Phe	Ala	Ala	Ile	Ala	Gly	Ala	Gly	Pro
	130					135					140				
Leu	Val	Gly	Pro	Val	Met	Ala	Ala	Gln	Met	Gly	Tyr	Leu	Pro	Gly	Thr
145					150					155					160
Leu	Trp	Ile	Ile	Leu	Gly	Val	Ile	Phe	Ala	Gly	Ala	Val	Gln	Asp	Tyr
				165					170					175	
Leu	Val	Leu	Trp	Val	Ser	Thr	Arg	Arg	Arg	Gly	Arg	Ser	Leu	Gly	Gln
			180					185					190		
Met	Val	Arg	Asp	Glu	Met	Gly	Thr	Val	Gly	Gly	Ala	Ala	Gly	Ile	Leu
		195					200					205			
Ala	Thr	Ile	Ser	Ile	Met	Ile	Ile	Ile	Ile	Ala	Val	Leu	Ala	Leu	Ile
	210					215					220				
Val	Val	Asn	Ala	Leu	Ala	Asp	Ser	Pro	Trp	Gly	Val	Phe	Ser	Ile	Thr
225					230					235					240
Met	Thr	Ile	Pro	Ile	Ala	Leu	Phe	Met	Gly	Val	Tyr	Leu	Arg	Tyr	Leu
				245					250					255	

```
Arg Pro Gly Arg Val Thr Glu Val Ser Ile Ile Gly Val Ala Leu Leu
            260                 265                 270

Leu Leu Ala Ile Val Ala Gly Gly Trp Val Ala Asp Thr Ser Trp Gly
        275                 280                 285

Val Glu Trp Phe Thr Trp Ser Lys Thr Thr Leu Ala Leu Ala Leu Ile
    290                 295                 300

Gly Tyr Gly Ile Met Ala Ala Ile Leu Pro Val Trp Leu Leu Leu Ala
305                 310                 315                 320

Pro Arg Asp Tyr Leu Ser Thr Phe Met Lys Ile Gly Val Ile Gly Leu
                325                 330                 335

Leu Ala Val Gly Ile Leu Phe Ala Arg Pro Glu Val Gln Met Pro Ser
            340                 345                 350

Val Thr Ser Phe Ala Leu Glu Gly Asn Gly Pro Val Phe Ser Gly Ser
        355                 360                 365

Leu Phe Pro Phe Leu Phe Ile Thr Ile Ala Cys Gly Ala Leu Ser Gly
    370                 375                 380

Phe His Ala Leu Ile Ser Ser Gly Thr Thr Pro Lys Leu Val Glu Lys
385                 390                 395                 400

Glu Ser Gln Met Arg Met Leu Gly Tyr Gly Gly Met Leu Met Glu Ser
                405                 410                 415

Phe Val Ala Met Met Ala Leu Ile Thr Ala Val Ile Leu Asp Arg His
            420                 425                 430

Leu Tyr Phe Ser Met Asn Ala Pro Leu Ala Leu Thr Gly Gly Asp Pro
        435                 440                 445

Ala Thr Ala Ala Glu Trp Val Asn Ser Ile Gly Leu Thr Gly Ala Asp
    450                 455                 460

Ile Thr Pro Glu Gln Leu Ser Glu Ala Ala Glu Ser Val Gly Glu Ser
465                 470                 475                 480

Thr Val Ile Ser Arg Thr Gly Gly Ala Pro Thr Leu Ala Phe Gly Met
                485                 490                 495

Ser Glu Ile Leu Ser Gly Phe Ile Gly Gly Ala Gly Met Lys Ala Phe
            500                 505                 510

Trp Tyr His Phe Ala Ile Met Phe Glu Ala Leu Phe Ile Leu Thr Thr
        515                 520                 525

Val Asp Ala Gly Thr Arg Val Ala Arg Phe Met Met Thr Asp Thr Leu
    530                 535                 540

Gly Asn Val Pro Gly Leu Arg Arg Phe Lys Asp Pro Ser Trp Thr Val
545                 550                 555                 560
```

```
Gly Asn Trp Ile Ser Thr Val Phe Val Cys Ala Leu Trp Gly Ala Ile
                565                 570                 575

Leu Leu Met Gly Val Thr Asp Pro Leu Gly Gly Ile Asn Val Leu Phe
            580                 585                 590

Pro Leu Phe Gly Ile Ala Asn Gln Leu Leu Ala Ala Ile Ala Leu Ala
        595                 600                 605

Leu Val Leu Val Val Val Val Lys Lys Gly Leu Tyr Lys Trp Ala Trp
    610                 615                 620

Ile Pro Ala Val Pro Leu Ala Trp Asp Leu Ile Val Thr Met Thr Ala
625                 630                 635                 640

Ser Trp Gln Lys Ile Phe His Ser Asp Pro Ala Ile Gly Tyr Trp Ala
                645                 650                 655

Gln Asn Ala Asn Phe Arg Asp Ala Lys Ser Gln Gly Leu Thr Glu Phe
            660                 665                 670

Gly Ala Ala Lys Ser Pro Glu Ala Ile Asp Ala Val Ile Arg Asn Thr
        675                 680                 685

Met Ile Gln Gly Ile Leu Ser Ile Leu Phe Ala Val Leu Val Leu Val
    690                 695                 700

Val Val Gly Ala Ala Ile Ala Val Cys Ile Lys Ser Ile Arg Ala Arg
705                 710                 715                 720

Ala Ala Gly Thr Pro Leu Glu Thr Thr Glu Glu Pro Asp Thr Glu Ser
                725                 730                 735

Glu Phe Phe Ala Pro Thr Gly Phe Leu Ala Ser Ser Arg Asp Lys Glu
            740                 745                 750

Val Gln Ala Met Trp Asp Glu Arg Tyr Pro Gly Gly Ala Pro Val Ser
        755                 760                 765

Ser Gly Gly His
    770
```

Patent claims

1. An isolated polynucleotide from coryneform bacteria, comprising a polynucleotide sequence which codes for the cstA gene, chosen from the group consisting of

   a) polynucleotide which is identical to the extent of at least 70% to a polynucleotide which codes for a polypeptide which comprises the amino acid sequence of SEQ ID No. 2,

   b) polynucleotide which codes for a polypeptide which comprises an amino acid sequence which is identical to the extent of at least 70% to the amino acid sequence of SEQ ID No. 2,

   c) polynucleotide which is complementary to the polynucleotides of a) or b), and

   d) polynucleotide comprising at least 15 successive nucleotides of the polynucleotide sequence of a), b) or c),

   the polypeptide preferably having the activity of carbon starvation protein A.

2. The polynucleotide as claimed in claim 1, wherein the polynucleotide is a preferably recombinant DNA which is capable of replication in coryneform bacteria.

3. The polynucleotide as claimed in claim 1, wherein the polynucleotide is an RNA.

4. The polynucleotide as claimed in claim 2, comprising the nucleic acid sequence as shown in SEQ ID No. 1.

5. The DNA as claimed in claim 2 which is capable of replication, comprising

   (i) the nucleotide sequence shown in SEQ ID No. 1, or

(ii) at least one sequence which corresponds to sequence (i) within the range of the degeneration of the genetic code, or

(iii) at least one sequence which hybridizes with the sequence complementary to sequence (i) or (ii), and optionally

(iv) sense mutations of neutral function in (i).

6. The polynucleotide sequence as claimed in claim 2, which codes for a polypeptide which comprises the amino acid sequences shown in SEQ ID No. 2.

7. Coryneform bacteria in which the cstA gene is enhanced, in particular over-expressed.

8. A process for the fermentative preparation of L-amino acids, in particular L-lysine, wherein the following steps are carried out:

a) fermentation of the coryneform bacteria which produce the desired L-amino acid and in which at least the cstA gene or nucleotide sequences which code for it are enhanced, in particular over-expressed;

b) concentration of the L-amino acid in the medium or in the cells of the bacteria, and

c) isolation of the L-amino acid.

9. The process as claimed in claim 8, wherein bacteria in which further genes of the biosynthesis pathway of the desired L-amino acid are additionally enhanced are employed.

10. The process as claimed in claim 8, wherein bacteria in which the metabolic pathways which reduce the formation of the desired L-amino acid are at least partly eliminated are employed.

11. The process as claimed in claim 8, wherein a strain transformed with a plasmid vector is employed, and the plasmid vector carries the nucleotide sequence which codes for the cstA gene.

12. The process as claimed in claim 8, wherein the expression of the polynucleotide which codes for the cstA gene is enhanced, in particular over-expressed.

13. The process as claimed in claim 8, wherein the regulatory properties of the polypeptide for which the polynucleotide cstA codes are increased.

14. The process as claimed in claim 8, wherein for the preparation of L-amino acids, in particular L-lysine, coryneform microorganisms in which at the same time one or more of the genes chosen from the group consisting of

14.1 the lysC gene which codes for a feed back resistant aspartate kinase,

14.2 the dapA gene which codes for dihydrodipicolinate synthase,

14.3 the gap gene which codes for glycerolaldehyde 3-phosphate dehydrogenase,

14.4 the pgk gene which codes for 3-phosphoglycerate kinase,

14.5 the pyc gene which codes for pyruvate carboxylase,

14.6 the tpi gene which codes for triose phosphate isomerase,

14.7 the lysE gene which codes for lysine export,

14.8 the zwa1 gene which codes for the Zwa1 protein,

is or are enhanced or over-expressed are fermented.

15. The process as claimed in claim 8, wherein for the preparation of L-amino acids, in particular L-lysine, coryneform microorganisms in which at the same time one or more of the genes chosen from the group consisting of

15.1 the pck gene which codes for phosphoenol pyruvate carboxykinase,

15.2 the pgi gene which codes for glucose 6-phosphate6 isomerase,

15.3 the poxB gene which codes for pyruvate oxidase,

15.4 the zwa2 gene which codes for the Zwa2 protein

is or are attenuated are fermented.

16. Coryneform bacteria which contain a vector which carries a polynucleotide as claimed in claim 1.

17. The process as claimed in one or more of the preceding claims, wherein microorganisms of the genus Corynebacterium glutamicum are employed.

18. A process for discovering RNA, cDNA and DNA in order to isolate nucleic acids, or polynucleotides or genes which code for carbon starvation protein A or have a high similarity with the sequence of the cstA gene, wherein the polynucleotide sequences as claimed in claim 1, 2, 3 or 4 are employed as hybridization probes.

19. The process as claimed in claim 18, wherein the hybridization is carried out under a stringency corresponding to at most 2x SSC.

New nucleotide sequences which code for the cstA gene

Abstract

The invention relates to an isolated polynucleotide comprising a polynucleotide sequence chosen from the group consisting of

a) polynucleotide which is identical to the extent of at least 70% to a polynucleotide which codes for a polypeptide which comprises the amino acid sequence of SEQ ID No. 2,

b) polynucleotide which codes for a polypeptide which comprises an amino acid sequence which is identical to the extent of at least 70% to the amino acid sequence of SEQ ID No. 2,

c) polynucleotide which is complementary to the polynucleotides of a) or b), and

d) polynucleotide comprising at least 15 successive nucleotides of the polynucleotide sequence of a), b) or c),

and a process for the fermentative preparation of L-amino acids using coryneform bacteria in which at least the cstA gene is present in enhanced form, and the use of the polynucleotide sequences as hybridization probes.